



## Estimation of lifetime reproductive success when reproductive status cannot always be assessed

Lauriane Rouan, Jean-Michel Gaillard, Yann Guédon, Roger Pradel

### ► To cite this version:

Lauriane Rouan, Jean-Michel Gaillard, Yann Guédon, Roger Pradel. Estimation of lifetime reproductive success when reproductive status cannot always be assessed. Thomson, D. L.; Cooch, E. G.; Conroy, M. J. Modeling Demographic Processes in Marked Populations, 3, Springer, pp.867-879, 2009, Environmental and Ecological Statistics, 978-0-387-78151-8. 10.1007/978-0-387-78151-8\_40 . hal-00831798

**HAL Id: hal-00831798**

**<https://inria.hal.science/hal-00831798>**

Submitted on 7 Jun 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Estimation of Lifetime Reproductive Success when reproductive status cannot always be assessed

L. Rouan, J-M. Gaillard, Y. Guédon, R. Pradel

**Abstract** The Lifetime Reproductive Success (LRS) of an individual i.e. the number of young raised during its lifespan is an indicator of its contribution to future generations and thus a measure of fitness. Nevertheless, the LRS is hard to estimate because of the difficulty to keep track of the outcome of each breeding attempt (successful or failed and, if successful, number of young raised). We propose two new methods to estimating the LRS that takes into account the uncertainty about the reproductive status when the individuals are not detected or when the reproductive status cannot be assessed. We illustrate these two methods using roe deer reproductive histories and discuss their advantages and disadvantages.

**Keywords:** Capture-Recapture, counting algorithm, fitness, generalized Viterbi algorithm, hidden Markov models

## 1 Introduction

Lifetime Reproductive Success (LRS) is a commonly used estimate of individual fitness (Clutton-Brock, 1988; Newton, 1989). It can be defined as the total number of offspring an individual produces over its entire lifespan after some critical stage has been successfully passed (e.g. number of weaned young in mammals or number of fledglings in birds, see Clutton-Brock (1988) for case studies). As

---

L. Rouan  
Centre d'Écologie Fonctionnelle et Évolutive, Montpellier, France  
e-mail: lauriane.rouan@cefe.cnrs.fr

J-M. Gaillard  
Université Claude Bernard, Lyon, France

Y. Guédon  
Centre de coopération Internationale en Recherche Agronomique pour le Développement,  
Montpellier, France

R. Pradel  
Centre d'Écologie Fonctionnelle et Évolutive, Montpellier, France

LRS is a measure of the lifetime reproductive performance of an individual, its actual calculation supposes the knowledge of the individual's entire reproductive history. Nevertheless, in wild populations, the exhaustive monitoring of a large number of individuals over a long time period is difficult; there are not only problems of detection but also problems in assessing the reproductive status. For instance, an individual seen during the breeding season could have produced and/or raised young or not and if it has, the number of its offspring could be difficult to determine. So, the capture-recapture (CR) data are inevitably incomplete. In the absence of established statistical tools, some "ad hoc" methods are used and consist of counting only the reproductive events observed leading thus to a 'minimum LRS' or of completing the reproductive histories by assuming that an individual not observed in a given year during the reproductive season, but captured (or observed) a previous year and captured (or observed) in a later year has some chance to be an effective breeder (Jensen et al., 2004). Pradel (2005) proposes a new approach, called multievent models, to deal with the uncertainty in the breeding states assessment. The specificity of these models, belonging to the hidden Markov models class (Rabiner, 1989; Ephraim and Merhav, 2002), lies in the distinction between the (hidden) states (in our case, the reproductive status) and the "events" observed and recorded in the reproductive histories (for example: 'not seen', 'seen with one offspring', 'seen during the breeding period', ...). Thus, contrary to the hypotheses of the now well established multistate CR models (Arnason, 1973), the underlying state of a captured individual is not necessarily known for certain i.e. each observation is potentially linked to several underlying states. Thus, a particular encounter history can be associated with different sequences of 'hidden states' (see Figure 1); but which of these has actually generated the sequence of observations (encounter history)? Answering this question would enable to calculate the LRS but it is, unfortunately, impossible. So, we suggest in this paper to estimate the LRS by taking into account all these potential sequences. We propose two different methods. The first method (Section 3.1) consists in the actual determination of the potential hidden state sequences in order of probability of occurrences ('associated probability'). It uses the Generalized Viterbi Algorithm (Foreman, 1993). To each sequence corresponds a number of young produced and/or raised and the LRS is estimated as the mean of these numbers weighted by the associated probabilities. But the calculation of the LRS does not require the knowledge of the timing of reproduction, only the number of reproductive episodes is required. So, we present a second method, less expensive in terms of computation time (section 3.2). It is an adaptation to the CR field of a counting algorithm initially designed for the general framework of the hidden semi-Markov chains (Guédon, 1999). It does not enable the access to the underlying states sequences but provides the distribution of the number of occurrences of any underlying state during the animal lifespan and enables us to estimate the LRS.

Both approaches make use of the common notation and are illustrated with a common example presented in the following section.

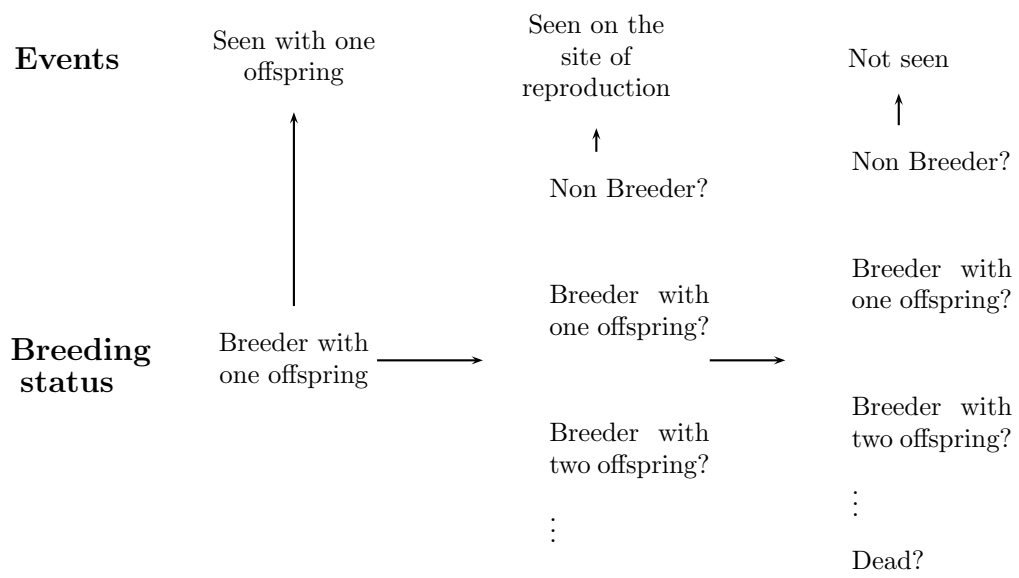


Figure 1: The events recorded at each capture occasion are potentially linked to several states. In other words different state sequences can generate the same encounter history.

## 2 Illustrating the two methods using the roe deer *Capreolus capreolus* life history

The 'multievent approach' is based on the differentiation of the observations made at each capture session and the actual reproductive status of the females. In other words, as illustrated in Figure 1 we consider two separate stochastic processes:

- the state process denoted  $(S_t)_{t \geq 0}$  that describes the succession of the reproductive statuses;
- the observation process denoted  $(O_t)_{t \geq 0}$  that describes the different 'events' observed over the study period.

These two processes are linked using what we call, hereafter, the 'observation probabilities' or 'conditional event probabilities'. Based on the roe deer life history, we can distinguish the following different 'events' (see e.g. Gaillard et al., 1998):

- 0: the animal is not seen;
- 1: the animal is seen without any fawn;
- 2: the animal is seen with one fawn;
- 3: the animal is seen with two fawns;
- 4: the animal is seen with three fawns.

The associated underlying states are:

- Non Breeder (NB);
- Breeder with one fawn (B1);
- Breeder with two fawns (B2);
- Breeder with three fawns (B3);
- Dead ( $\dagger$ ).

This defines the event and state sets:  $\Omega = \{0, 1, 2, 3, 4\}$  and  $E = \{NB, B1, B2, B3, \dagger\}$ . It remains to specify the parameters that define the multievent model and appear in the different algorithms that we propose to estimate the LRS. There are three kinds of parameters:

1. the transition probabilities:
  - $\phi_{ij}^t = P(S_{t+1} = j | S_t = i)$ , probability of being in state  $j$  at  $t + 1$  for an animal in state  $i$  at  $t$ ;
2. the initial state probabilities

- $\pi_i^t = P(S_t = i)$ , probability of being in state  $i$  when first captured at  $t$ . In the general case, females are initially marked during early life to be considered as known-aged individuals. For instance, roe deer are often marked in their first winter at about 8 months of age and can only be in state 'non breeder' since the first reproduction is only possible at 2-year-old (Gaillard et al., 1992). Thus  $\pi_{NB}^t = 1$ .

3. the conditional event probabilities:

- $b_s^t(o) = P(O_t = o | S_t = s)$ , probability of event  $o$  conditional on current state  $s$ .

These last parameters can actually be expressed as functions of the familiar encounter probabilities of the multistate models:

- $p_j^t$ , probability to be encountered in site  $j$  at time  $t$  for an individual alive and in site  $j$  at that time.

For roe deer, we can assume that the uncertainty in assessing the reproductive status is only due to problems of detection i.e. if a female is seen, its reproductive status corresponds exactly to the observation made in the right period (September-December) because fawns closely follow their mother during that period (see Gaillard et al., 2000). So, most of the  $b_s^t(o)$  are equal to zero. The only not null parameters are:  $b_{NB}^t(1) = p_{NB}^t$ ,  $b_{B1}^t(2) = p_{B1}^t$ ,  $b_{B2}^t(3) = p_{B2}^t$ ,  $b_{B3}^t(4) = p_{B3}^t$ ,  $b_j^t(0) = 1 - p_j^t$  for  $j \in \{NB, B1, B2, B3\}$  and  $b_{\dagger}^t(0) = 1$ . This assumption has been made in agreement with our knowledge about roe deer life history in order to simplify the statistical model but is not necessary for any other study. In the general case, two kinds of uncertainty can occur: the lack of detection at some reproductive attempts and an uncertain determination for animals detected. Both kinds can be easily accounted for in the two methods detailed below.

The estimation of these different parameters is the first step, common to our two approaches, to estimate the LRS.

### 3 Estimation of the Lifetime Reproductive Success

To calculate the LRS, one needs to know the number of offspring an individual has produced at each reproductive occasion. The first, intuitive, idea is to reconstruct the reproductive life of the individual.

### 3.1 Method 1: estimating LRS using the generalized Viterbi algorithm

#### 3.1.1 Method

Given any observation sequence  $O = o_1 o_2 \dots o_T$ , the generalized Viterbi algorithm (Foreman, 1993) seeks to find the  $L$  ( $L \geq 1$ ) most probable underlying state sequences  $\{S^i = s_1^i s_2^i \dots s_T^i / i = 1, 2, \dots, L\}$  maximizing the associated probability  $P(S|O)$ . This algorithm has been originally designed for the area of automatic speech recognition and is adapted to the treatment of homogeneous Markov chains and observations sequences starting at time  $t = 1$ . Its use for the CR data needs some specific adaptations to:

1. allow for time varying parameters;
2. handle histories starting after date 1;
3. use a stopping criteria based on the cumulative probability of occurrence i.e. our 'adapted' generalized Viterbi algorithm doesn't generate a fixed number of state sequences but more precisely generates the  $L_\alpha$  state sequences such that the cumulative probability  $\sum_{i=1}^{L_\alpha} P(S^i|O)$  reaches a fixed threshold  $\alpha$ . In this way, to obtain all the possible state sequences,  $\alpha$  must be fixed to 1.

#### 3.1.2 LRS estimation

The parameters of the retained model, estimated using program E-SURGE (Choquet et al., 2007), are used to calculate the different quantities needed to define the Viterbi algorithm. The second step of this method is to apply the algorithm to a particular encounter history  $O$  thus generating the  $L_\alpha$  state sequences and their associated probability. Finally the LRS is estimated as the weighted mean of the number of offspring given by the different state sequences:

$$LRS = \frac{\sum_{i=1}^{L_\alpha} n_{offspring}^i P(S^i|O)}{\sum_{i=1}^{L_\alpha} P(S^i|O)}$$

#### 3.1.3 Examples

Let us illustrate this approach using two encounter histories taken from the data set

1.  $O^1 = 0000000000000000000010402202220$
2.  $O^2 = 0000000000001323203222100000000$

the second one being 'less complete' than the first in the sense that there are fewer detection issues (number of 0 after the individual first capture) in  $O^1$  than in  $O^2$ .

To illustrate this first method, we have chosen to fix  $\alpha$  to one.

The application of our 'adapted' generalized Viterbi algorithm to  $O^1$  provides the nine following state sequences (in descending order of associated probability):

- $S^1 = \text{NB NB B3 B1 B1 B1 B1 B1 B1 B1 B1}$  with  $P(S^1|O^1) \approx 0.2228$ ;
- $S^2 = \text{NB B1 B3 B1 B1 B1 B1 B1 B1 B1 B1}$  with  $P(S^2|O^1) \approx 0.1827$ ;
- $S^3 = \text{NB B3 B3 B1 B1 B1 B1 B1 B1 B1 B1}$  with  $P(S^3|O^1) \approx 0.1421$ ;
- $S^4 = \text{NB NB B3 B1 B1 B1 B1 B1 B1 B1 } \dagger$  with  $P(S^4|O^1) \approx 0.1359$ ;
- $S^5 = \text{NB B1 B3 B1 B1 B1 B1 B1 B1 B1 } \dagger$  with  $P(S^5|O^1) \approx 0.1114$ ;
- $S^6 = \text{NB B3 B3 B1 B1 B1 B1 B1 B1 B1 } \dagger$  with  $P(S^6|O^1) \approx 0.0866$ ;
- $S^7 = \text{NB NB B3 B1 B1 B1 B1 B1 B1 B1 B2}$  with  $P(S^7|O^1) \approx 0.0482$ ;
- $S^8 = \text{NB B1 B3 B1 B1 B1 B1 B1 B1 B1 B2}$  with  $P(S^8|O^1) \approx 0.0395$ ;
- $S^9 = \text{NB B3 B3 B1 B1 B1 B1 B1 B1 B1 B2}$  with  $P(S^9|O^1) \approx 0.0307$ .

Thus, the LRS is estimated as:

$$\begin{aligned} LRS_1 &= 11 \times 0.2228 + 12 \times 0.1827 + 14 \times 0.1421 + 10 \times 0.1359 + 11 \times 0.1114 \\ &\quad + 13 \times 0.0866 + 12 \times 0.0482 + 13 \times 0.0395 + 15 \times 0.0307 \\ &\approx 11.8965 \end{aligned}$$

The same method applied to the encounter history  $O^2$  generates 492 state sequences! The 'best' and 'worst' sequences with their associated probability are:

- $S^1 = \text{NB B2 B1 B2 B1 B1 B2 B1 B1 B1 NB } \dagger \dagger \dagger \dagger \dagger \dagger \dagger \dagger$  with  $P(S^1|O^2) \approx 0.6991$ .
- $S^{492} = \text{NB B2 B1 B2 B1 B1 B2 B1 B1 B1 NB B2 B1 B2 B1 B2 B2 B1 B2}$  with  $P(S^{492}|O^2) \approx 6.1806e - 007$ .

The estimated LRS is, for this second encounter history: 12.7639.

The generalized Viterbi algorithm is thus a reasonable approach to estimate the LRS but presents some limits in terms of computation time. The number of reconstructed state sequences is variable depending on the associated encounter history; the computation time can vary between few seconds like in case of  $O^1$  and tens of minutes for  $O^2$ . However, as the mere calculation of the LRS does not require the knowledge of the exact sequence of reproductive events but only the total number of offspring produced and/or raised, we propose another approach for estimating the LRS, which yields the same estimate as the generalized Viterbi algorithm with a threshold  $\alpha = 1$  but is much less expensive in computation time.



## 3.2 Method 2: estimating LRS using the counting algorithm

### 3.2.1 Method

Our general idea to calculate the LRS is to count the occurrences of the breeding hidden states. More precisely, this counting algorithm determines, giving an encounter history  $O = o_e o_{e+1} \cdots o_T$ , the probability  $P(N_k(T) = n | o_e o_{e+1} \cdots o_T)$  that an underlying state  $k$  has occurred  $n$  times ( $n \in \{0, 1, 2, \dots, T - e + 2\}$ ,  $e$  is the date of first capture and  $T$  the last capture occasion) by the end of the study.

Let us so introduce the auxiliary quantities

$$\beta_t(i, n) = P(S_t = i, N_k(t) = n, O_e = o_e, \dots, O_t = o_t) \quad (1)$$

The quantity  $\beta_t(i, n)$  is the unconditional probability that the underlying state  $k$  has occurred  $n$  times until the date  $t$  and that the underlying state at this date is  $i$ .

The probability  $P(N_k(T) = n | o_e o_{e+1} \cdots o_T)$  that the state  $k$  occurs  $n$  times can be expressed using these quantities as follows :

$$P(N_k(T) = n | o_e o_{e+1} \cdots o_T) = \frac{\sum_{i=1}^N \beta_T(i, n)}{\sum_{n=0}^T \sum_{i=1}^N \beta_T(i, n)} \quad (2)$$

The denominator corresponds to the conditioning on the encounter history  $O$ . The counting algorithm corresponds then to the recursive calculation of  $\beta_t(i, n)$  for all  $t$  in  $\{0, 1, 2, \dots, T\}$  and  $n$  in  $\{0, 1, 2, \dots, t\}$  (see Appendix A) .

### 3.2.2 LRS estimation

As in the Viterbi algorithm case, the first step is to introduce the estimated parameters of the retained model in the algorithm. In a second step, applying it to an encounter history  $O = o_e o_{e+1} \cdots o_T$ , provides for each underlying state  $k$  the probabilities:  $P(N_k(T) = n | O)$  for  $n$  in  $\{0, 1, 2, \dots, T\}$  i.e. the distribution of the number of occurrences of each state  $k$ .

For each state  $k$ , the conditionnal expectancy  $E[N_k(T) | O]$  can be interpreted as the average number of occasions in state  $k$ . We can easily show that

$\sum_{\text{all states } k} E[N_k(T) | O] = T - e + 1$  is the number of capture occasions since the first capture of the individual (see Appendix B). Given these quantities, the LRS is estimated in this way:

$LRS = \sum_{k=1}^K k \times E[N_{Bk}(T) | O]$  where  $K$  is the maximum number of offspring an individual can produce at each breeding season and  $Bk$  corresponds to the status "breeder with  $k$  offspring".

### 3.2.3 Examples

In our examples, only the states B1, B2 and B3 are interesting, so we will give only the distributions relative to these particular states.

For the first encounter history  $O^1$  of section 3.1.3:

$n$	0	1	2	3	4	5	6	7
$P(N_{B1}(T) = n O^1)$	0	0	0	0	0	0	0	0.3015
$P(N_{B2}(T) = n O^1)$	0.8815	0.1185	0	0	0	0	0	0
$P(N_{B3}(T) = n O^1)$	0	0.7405	0.2595	0	0	0	0	0

$n$	8	9	10	11
$P(N_{B1}(T) = n O^1)$	0.5158	0.1827	0	0
$P(N_{B2}(T) = n O^1)$	0	0	0	0
$P(N_{B3}(T) = n O^1)$	0	0	0	0

The estimated LRS is so:

$$\begin{aligned}
LRS_1 &= 1 \times E[N_{B1}(T)|O^1] + 2 \times E[N_{B2}(T)|O^1] + 3 \times E[N_{B3}(T)|O^1] \\
&= 1 \times 7.8811 + 2 \times 0.1185 + 3 \times 1.2595 \\
&\approx 11.8966
\end{aligned}$$

For the second encounter history  $O^2$ :

$n$	0	1	2	3	4	5	6
$P(N_{B1}(T) = n O^2)$	0	0	0	0	0	0	0.7237
$P(N_{B2}(T) = n O^2)$	0	0	0	0.9048	0.0729	0.0151	0.0049
$P(N_{B3}(T) = n O^2)$	0.9971	0.0029	0	0	0	0	0

$n$	7	8	9	10	11	12
$P(N_{B1}(T) = n O^2)$	0.1197	0.1095	0.0329	0.0099	0.0030	0.0009
$P(N_{B2}(T) = n O^2)$	0.0016	0.0005	0.0002	0.0001	0	0
$P(N_{B3}(T) = n O^2)$	0	0	0	0	0	0

$n$	13	14	15	16	17	18	19
$P(N_{B1}(T) = n O^2)$	0.0002	0	0	0	0	0	0
$P(N_{B2}(T) = n O^2)$	0	0	0	0	0	0	0
$P(N_{B3}(T) = n O^2)$	0	0	0	0	0	0	0

The estimated LRS is so:

$$\begin{aligned}
LRS_2 &= 1 \times 6.4998 + 2 \times 3.1282 + 3 \times 0.0029 \\
&\approx 12.7649
\end{aligned}$$

In the two cases, the counting algorithm provides the same results as the generalized Viterbi algorithm but with a computation time approaching few seconds.

## 4 Discussion

This article has presented new general methods to estimate Lifetime Reproductive Success with missing values for some reproductive events (generated either by a lack of detection or by an uncertain reproductive status). The two methods provided identical LRS estimates. The main difference between the two methods lies in the reconstruction or not of the possible reproductive sequences associated with a given encounter history. The 'Viterbi approach' is the more detailed in that it reveals the state sequences associated with the encounter history and their probability of occurrence. The LRS is then simply calculated as the weighted average of the numbers of offspring provided by these state sequences. However, this algorithm can be very expensive in terms of computation time depending on the level of uncertainty of the encounter history. The counting algorithm only provides the expected distribution of the number of offspring produced and/or raised over the lifetime. The exact sequence of reproductive events along the lifetime is not informed using this latter method but is much faster than the Viterbi algorithm. The counting algorithm seems thus the most appropriate method to estimate the LRS. However, LRS is just one measure of single-generation individual fitness, and other single-generation measurements have been proposed to be better in some situations. For instance, the 'individual  $\lambda$ ' method (McGraw and Caswell, 1996) accounts for the timing of reproductive events, which LRS does not (Käär and Jokela, 1998). Although a recent comparative analysis of fitness measures showed that LRS generally performs better than 'individual  $\lambda$ ' (Brommer et al., 2004), the latter measure is likely to be preferred in markedly increasing or decreasing populations when the timing of reproductive events has a major impact on fitness. By using the Viterbi algorithm, biologists could not only obtain an estimate of LRS, but also get estimates of reproductive output at each breeding attempt. Therefore, the calculation of individual  $\lambda$ , or of the promising measure of individual contributions (Coulson et al., 2006) would be straightforward by using the Viterbi algorithm. On the other hand, only an estimate of LRS can be obtained using the counting method.

In conclusion, the counting method we proposed here should allow biologists to get a way of estimating quickly LRS in a large range of field conditions. By using the Viterbi method, biologists could even obtain different time-sensitive and time-insensitive fitness measures. These two methods are currently implemented in MATLAB and will be soon available in program E-SURGE.

## 5 Appendices

### 5.1 Appendix A: the counting algorithm

As mentioned above for the adaptation of the generalized Viterbi algorithm, in capture-recapture studies, probabilities are usually calculated conditionally on the first capture of the individuals ; so the recursion will begin with  $t = e$ .

### 1. Initialization

$$\begin{aligned} t &= e, \\ i &\in E, \\ n &\in \{0, 1\}, \end{aligned}$$

If  $i = k$ ,

$$\begin{aligned} \beta_e(k, 0) &= 0 \\ \beta_e(k, 1) &= P(S_e = k, N_k(e) = 1, O_e = o_e) \\ &= P(S_e = k, O_e = o_e) \\ &= \pi_k^e b_k^e(o_e) \end{aligned}$$

Si  $i \neq k$ ,

$$\begin{aligned} \beta_e(i, 0) &= P(S_e = i, N_k(e) = 0, O_e = o_e) \\ &= P(S_e = i, O_e = o_e) \\ &= \pi_i^e b_i^e(o_e) \\ \beta_e(k, 1) &= 0 \end{aligned}$$

### 2. Recursion

$$\begin{aligned} t &\in \{e + 1, e + 2, \dots, T\}, \\ i &\in E, \\ n &\in \{0, 1, 2, \dots, t\}, \end{aligned}$$

If  $i = k$ ,

$$\begin{aligned} \beta_t(i, n) &= P(S_t = k, N_k(t) = n, O_e = o_e, \dots, O_t = o_t) \\ &= \sum_{j=1}^N P(S_t = k, S_{t-1} = j, N_k(t-1) = n-1, O_e = o_e, \dots, O_{t-1} = o_{t-1}, O_t = o_t) \end{aligned}$$

If  $i \neq k$ ,

$$\begin{aligned} \beta_t(i, n) &= P(S_t = i, N_k(t) = n, O_e = o_e, \dots, O_t = o_t) \\ &= \sum_{j=1}^N P(S_t = i, S_{t-1} = j, N_k(t-1) = n, O_e = o_e, \dots, O_{t-1} = o_{t-1}, O_t = o_t) \end{aligned}$$

But, for all time  $t$ , all state  $\lambda$  and all number  $m$  :

$$\begin{aligned} P(S_t = \lambda, S_{t-1} = j, N_k(t-1) = m, O_e = o_e, \dots, O_{t-1} = o_{t-1}, O_t = o_t) \\ = P(O_t = o_t | S_t = \lambda, S_{t-1} = j, N_k(t-1) = m, O_e = o_e, \dots, O_{t-1} = o_{t-1}) \end{aligned}$$

$$\begin{aligned}
& \times P(S_t = \lambda, S_{t-1} = j, N_k(t-1) = m, O_e = o_e, \dots, O_{t-1} = o_{t-1}) \\
= & P(O_t = o_t | S_t = \lambda) \times P(S_t = \lambda | S_{t-1} = j, N_k(t-1) = m, O_e = o_e, \dots, O_{t-1} = o_{t-1}) \\
& \times P(S_{t-1} = j, N_k(t-1) = m, O_e = o_e, \dots, O_{t-1} = o_{t-1}) \\
= & P(O_t = o_t | S_t = \lambda) \times P(S_t = \lambda | S_{t-1} = j) \\
& \times P(S_{t-1} = j, N_k(t-1) = m, O_e = o_e, \dots, O_{t-1} = o_{t-1})
\end{aligned}$$

So :

If  $i = k$ ,

$$\begin{aligned}
\beta_t(i, n) &= \sum_{j=1}^N b_k^t(o_t) \phi_{jk}^{t-1} \beta_t(j, n-1) \mathbf{1}_{[n>0]} \\
&= b_k^t(o_t) \sum_{j=1}^N \beta_{t-1}(j, n-1) \phi_{jk}^{t-1} \mathbf{1}_{[n>0]}
\end{aligned}$$

If  $i \neq k$ ,

$$\begin{aligned}
\beta_t(i, n) &= \sum_{j=1}^N b_i^t(o_t) \phi_{ji}^{t-1} \beta_t(j, n) \\
&= b_i^t(o_t) \sum_{j=1}^N \beta_{t-1}(j, n) \phi_{ji}^{t-1}
\end{aligned}$$

where

$$\begin{cases} \mathbf{1}_{[n>0]} = 0, & \text{if } n \leq 0 \text{ (here, if } n = 0) ; \\ \mathbf{1}_{[n>0]} = 1, & \text{if } n > 0. \end{cases}$$

## 5.2 Appendix B: $\sum E_{all \text{ states } k} [N_k(T)|O]$ = number of occasions

For any state  $k$  and encounter history  $O = o_e o_{e+1} o_{e+2} \dots o_T$ ,

$$\begin{aligned}
E[N_k(T)|O] &= E \left[ \sum_{t=e}^T \mathbf{1}_{[S_t=k]} | O \right] \\
&= \sum_{t=e}^T E[\mathbf{1}_{[S_t=k]} | O] \\
&= \sum_{t=e}^T P(S_t = k | O)
\end{aligned}$$

Thus,

$$\sum_{all \text{ states } k} E[N_k(T)|O] = \sum_{all \text{ states } k} \sum_{t=e}^T P(S_t = k | O)$$

$$\begin{aligned}
&= \sum_{t=e}^T \sum_{\text{all states } k} P(S_t = k|O) \\
&= \sum_{t=e}^T 1 \\
&= T - e + 1
\end{aligned}$$

## References

- Arnason, A. N. (1973). The estimation of population size, migration rates and survival in a stratified population. *Researches on Population Ecology*, 15(1):1–8.
- Brommer, J. E., Gustafson, L., Pietiainen, H., and Merila, J. (2004). Single-generation estimates of individual fitness as proxies for long-term genetic contribution. *American Naturalist*, 163:505–517.
- Choquet, R., Rouan, L., and Pradel, R. (2007). Program E-SURGE: a software application for fitting multievent models. *Environmental and Ecological Statistics*, in revision.
- Clutton-Brock, T. H. (1988). *Reproductive Success. Studies of Individual Variation in Contrasting Breeding Systems*. University of Chicago Press, Chicago.
- Coulson, T., Benton, T. G., Lundberg, P., Dall, S. R. X., Kendall, B. E., and Gaillard, J. M. (2006). Estimating individual contributions to population growth: evolutionary fitness in ecological time. *Proceedings of the Royal Society of London, Series B*, 273:547–555.
- Ephraim, Y. and Merhav, N. (2002). Hidden markov processes. *IEEE Transactions on Information Theory*, 48(6):1518–1569.
- Foreman, L. A. (1993). Generalization of the viterbi algorithm. *Journal of Mathematics Applied in Business and Industry*, 4:351–367.
- Gaillard, J. M., Andersen, R., Delorme, D., and Linnell, J. (1998). Family effects on growth and survival of juvenile roe deer. *Ecology*, 79:2878–2889.
- Gaillard, J. M., Festa-Bianchet, M., Delorme, D., and Jorgenson, J. (2000). Body mass and individual fitness in female ungulates: bigger is not always better. *Proceedings of the Royal Society of London, Series B*, 267:471–477.
- Gaillard, J. M., Sempere, A. J., Boutin, J. M., Van Laere, G., and Boisaubert, B. (1992). Effects of age and body weight on the proportion of females breeding in a population of roe deer (*Capreolus capreolus*). *Canadian Journal of Zoology*, 70:1541–1545.
- Guedon, Y. (1999). Computational methods for discrete hidden semi-markov chains. *Applied Stochastic Models in Business and Industry*, 15:195–224.

- Jensen, H., Saether, B. E., Ringsby, T. H., Tufto, J., Griffith, S. C., and Ellegren, H. (2004). Lifetime reproductive success in relation to morphology in the house sparrow *passer domesticus*. *Journal of Animal Ecology*, 73:599–611.
- Käär, P. and Jokela, J. (1998). Natural selection on age-specific fertilities in human females: comparison of individual-level fitness measures. *Proceedings of the Royal Society of London, Series B*, 265:2415–2420.
- McGraw, J. B. and Caswell, H. (1996). Estimation of individual fitness from life-history data. *American Naturalist*, 147:47–64.
- Newton, I. (1989). *Lifetime Reproduction in Birds*. Academic Press, London.
- Pradel, R. (2005). Multievent : An extension of multistate capture-recapture models to uncertain states. *Biometrics*, 61(2):442–447.
- Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77:257–286.